

Универсален Изкуствен Разум

Двигатели на човешкото поведение. Бихевиоризъм - учене с подсилване. Учение с учител и подражание.

4/2010

Двигател на човешкото поведение?

- Секс?
- Алчност?
- Жажда за власт?
- Любов
- Алтруизъм
- Вяра
- Стремение към съвършенство
- ...
-

Двигател на човешкото поведение?

- Какво всъщност са алчността, секса, жаждата за власт?
- Какво е любовта? Защо търсим любов?
- Какво е алтруизмът?
- Нерационални ли са любовта и вярата?
- **Има ли първичните елементи, с които могат да се изведат и обяснят всички чувства и мотиви на поведението?**
- **Как се придобиват нови поведения и защо?**

Какво е поведение?

- Начин по който някой реагира в дадена ситуация.
- Начин по който някой постига цели.
- Съвкупност/последователност от действия, схеми, похвати за постигане на промени в средата и в себе си.
- Неща, които някой прави.

Йерархия на нуждите на Maslow



*Прекалено високо ниво на описание

Maslow's Hierarchy of Needs

Abraham Maslow, Motivation and Personality, 2nd ed., Harper & Row, 1970

http://www.union.umd.edu/GH/basic_needs/index.htm

Как се дресират животни?

- Птици и бозайници
- Принцип на тоягата и моркова



- *Влечуги – прекалено примитивен краен мозък*

Октоподи

- Типичните мекотели – около 20 хил. дифузно разположени неврони
- **Октопод – 300 милиона неврони в мозък с дялове.** Интелигентност, сравнима с птици и дори бозайници.
- Учи се и играе - разпознава геом. фигури и когато са завъртяни, решава проблеми, изучава и помни лабиринти

<http://www.youtube.com/watch?v=EORp9mR2iOA&feature=related>

http://www.youtube.com/watch?v=jN99Kx_ghC8&NR=1

<http://www.youtube.com/watch?v=mFvxZet0DfQ&feature=related>

<http://www.youtube.com/watch?v=VLLQOK1gZE4&feature=related>

<http://www.youtube.com/watch?v=MWoUI1dqzLc&feature=related>

<http://www.slate.com/id/219221>

Дресиране на... хора?

- Мозъкът на хората не е по-различен от мозъка на другите бозайници в начина по който се учи.
- Защо бебето обича и слуша майка си, а не някой непознат?
- Защо човек с посттравматичен синдром не иска нищо да му напомня за травматичната ситуация?
- Защо когато човек открие нещо, което му е забавно и приятно, продължава да го прави?
- Защо когато някой човек ви е неприятен, го отбягвате, а когато ви е приятен, ви привлича и го търсите?

Дресиране на... хора?

- Защо хората правят абсурдни и безсмислени неща, като да се завъртят три пъти, да плюят и да хвърлят камък, след като са видели черна котка?
- Защо хората спазват ритуали?
- Защо хората се срамуват от абсурдни неща? (почти гола жена, прикрива само зърната си с ръце, виждането им „я обезчестява“)
- Защо „лъжицата трябва да се постави отдясно, вилицата отляво, ножът ...“, за да се покажат „добри обноски“.
- Защо когато звучи химнът трябва се изправите?

Как се дресират хора?

- Принцип на „тоягата и моркова“ ...

НАГРАДИ и НАКАЗАНИЯ

в подходящи моменти, ситуации и количества
...моделират поведението ни в бъдеще...

Награди и наказания

- **Образование** - оценки, поощрения, класове
- **Конкурси и състезания** - класиране, поощрения, награди
- **Морал** - порицание, похвала, уважение, презрение
- **Закони** - глоби, затвор, данъчни облекчения
- **Социално положение** – привилегии, власт, титли, пари
- **Служебно положение** – привилегии, заплата, власт
- **Войска** – повишение, понижение, ордени, разжалване
- **Семейство** – бой, отнемане на джобни, забрани; ласка, подаръци, позволения; и т.н.

Класическо кондициониране

- Кучето на Павлов
- Храна --> слюноотделяне
- (Лампа или Звънец) + Храна --> слюноотделяне
- Лампа или Звънец --> слюноотделяне
- **Повтарящи се времево близки стимули, един от които значим, се свързват в условен рефлекс.**
- **Кондициониране**
(от condition – условие)
Асоциативно обучение



Класическо кондициониране

- **Безусловен стимул*** – миризма на храна, рефлекс - слюноотделяне
- **Условен стимул** – звънец
- Страхова невроза при деца, които в ранна възраст са имали травматично преживяване с животни, вода, при което много са се уплашили.

* Възможно е слюноотделянето от миризма технически също да е условен рефлекс – не знам дали новороденото куче веднага отделя слюнка при миризма на храна; но, когато се е хранило първите пъти (поставяло е храна в устата си, това безусловно предизвиква слюнка), да се е получил този рефлекс.

Теории за Клас. Кондициониране (КК)

- **Стимул-Реакция (S-R)** – механично, без осмисляне. Кучето свързва звънеца със слюноотделянето.
- **Стимул-Стимул (S-S)** – когнитивна обработка; звънецът кара кучето да **очаква храна**.

Плъх – замръзва на място и се ослушва при:

- Безусловен стимул – силен звук (опасност)
- Успоредно - условен стимул – **лампа**
- Не се случва нищо лошо на плъха.

Светването само на лампата не предизвиква реакцията на безусловния стимул - плъхът знае че лампата не е знак за опасност. Доказателство за теорията SS.

„Портокал с часовников механизъм“

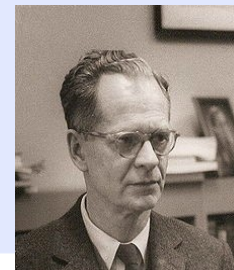


<http://video.google.com/videoplay?docid=-7656347491310375572#>

<http://research.twenkid.com>

Тодор Арнаудов

Оперантно кондициониране

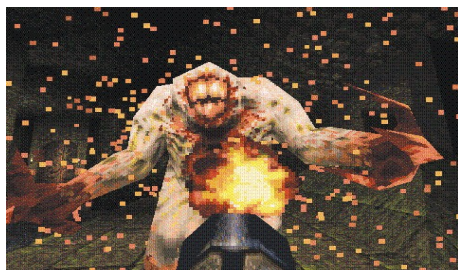


- Промяна/настройка на волевото поведение.
- Скинър - опити с гълъби и плъхове, и лостче.
- Взаимодействие със средата.
- Влияе се от последствията (бъдещето), а не от предишни стимули (при КК).
- **Бихевиоризъм** – дял от психологията, наука за поведението на животните и хората (агенти), която използва научен подход за измерване и промяна на поведението в желана посока и граници.

Видеоигри



- Неписани правила за разработка на игри:
 - награди и наказания
 - нарастваща трудност/избор на трудност
 - търсене на награди, бонуси; избягване на наказания



Награди/Наказания	Жанр игри
Брой/номер (ЧИСЛА): точки, фрагове, животи, ниво, победи, място в класиране, време и т.н.	Jump & Run, Puzzle, 3D Shooter, Fighting, RPG, Racing, Sport, Strategical, Quest, ...

Игри с коли

Награди:

position--

time-- (за обиколка)

damage--

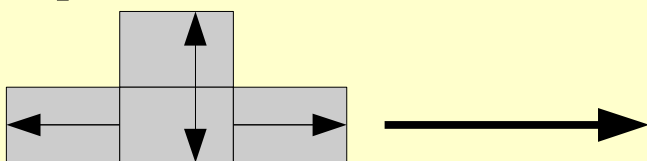
Наказания:

position++

time++

damage++

Управление



Състояние



Взаимовръзки (корелации)

Скорост++ --> Време--

Повреда++ --> Скорост --

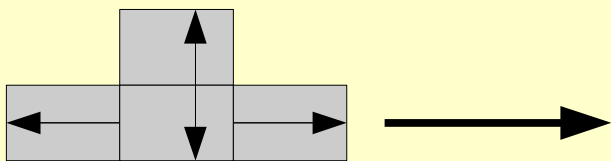
Скорост-- --> Място++

- Различаване на **права** и **завой**

- Ъгъл на влизане и скорост в завоя – скорост на излизане от завоя

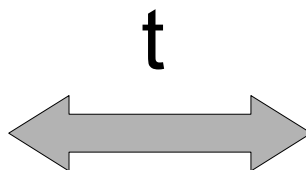
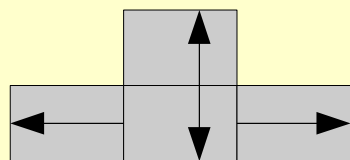
- Оптимална скорост и ъгъл за всеки завой

Управление

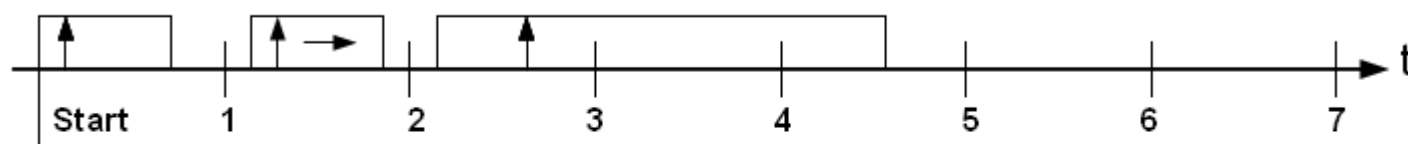


Оптимизация

Управление



Скорост,
Време,
Повреда,
Място



Кое действие да се избере в даден момент, за да се максимизира общата награда и да се минимизира наказанието.



Оперантно кондициониране

- Откриване на действията, които **увеличават вероятността** за постигане до награда
- Проби, експерименти и запомняне/натрупване на наблюдаваното развитие на събитията.
- В игрите с коли – ако дам повече газ на този завой, ако вляза по-остро, ако натисна по-късно спирачка – как се променя наградата?*
- **Опитът** позволява да **предсказваме**

** не разглеждаме проблема с влиянието на други коли/играчи – баланс скорост/риск от катастрофа*

Когнитивни задачи за животни

- Когато животното извърши желаното действие – получава награда.
- Внимателно животното се насочва да върши желаните действия и се награждава, ако ги извърши успешно. То запомня последователностите.
- **Така се изследват умствените способности на животни.**

Разпознаване на геометрични фигури (дори с октоподи), запомняне на изображения (гълъби), ориентиране в лабиринти (плъхове, мишки) и др.



Шимпанзетата имат фотографска памет

<http://www.youtube.com/watch?v=zJAH4>

Компоненти на оперантното кондициониране

- **Подсилване (*reinforcement*)** – последица на поведението (стимул – храна, звук, лампичка ...), която кара поведението да се случва по-често.
- **Наказание (*punishment*)**– последица на поведението, която кара поведението да се случва по-рядко.
- **Затихване (*extinction*)** – поведението няма последица - нито положителна, нито отрицателна. Честотата намалява и евентуално поведението се загубва (забравя).

Подсилване

- Положително подсилване (подсилване) – когато поведението се следва от приятен стимул, който увеличава вероятността от това поведение. *Плъх – натискане на ръчка, за да падне парченце храна.*
- Отрицателно подсилване (отбягване) – поведението се следва от премахване на неприятен стимул, което също увеличава вероятността от това поведение. *Плъх – ръчката спира звученето на силен звук.*

Наказване

- Положително наказване (наказване) – поведението се следва от неприятен стимул, който намалява честотата на поведението.
Плъх – натиска ръчката - удар или силен звук.
- Отрицателно наказване („глобяване“) – поведението се следва от *премахване на приятен стимул*, което също намалява честотата на поведението.
Бебе – удря другарче – взимате му близалката...

Учене с подсилване

- Използване на механизмите и принципите на оперантното кондициониране за учене/придобиване на нови поведения.
- **Метод в машинното обучение (по-късна лекция) - *Reinforcement Learning*.**
- *Утилитаризъм ~ Бихевиоризъм ~ Учене с подсилване* - в обществото и икономиката - стремеж към максимизиране на печалбата и минимизиране на загубите и търсене на поведения, които постигат тази цел.

Ефективност на наградите и наказанията за промяната в поведението

- **Задоволеност** - *сита котка не гони мишки*
- **Времева близост между поведението и последствията** – ако детенце счупи ваза и майка му го плесне, то ще разбере по-ясно, отколкото ако след една седмица му обясни, че е лошо да се чупят вази.
- **Устойчивост** – наградите или наказанията трябва да се появяват устойчиво и систематично, за да повлияят на поведението.
- **Съотношение между награда и наказание (риск)** – *цената на фиша струва ли си възможната печалба от лотарията.*

Учене с подсилване

- **Универсален метод за самообучение в произволна среда.**
- Системата/агентът трябва да може да познае и измери кога получава награда и кога е наказан!

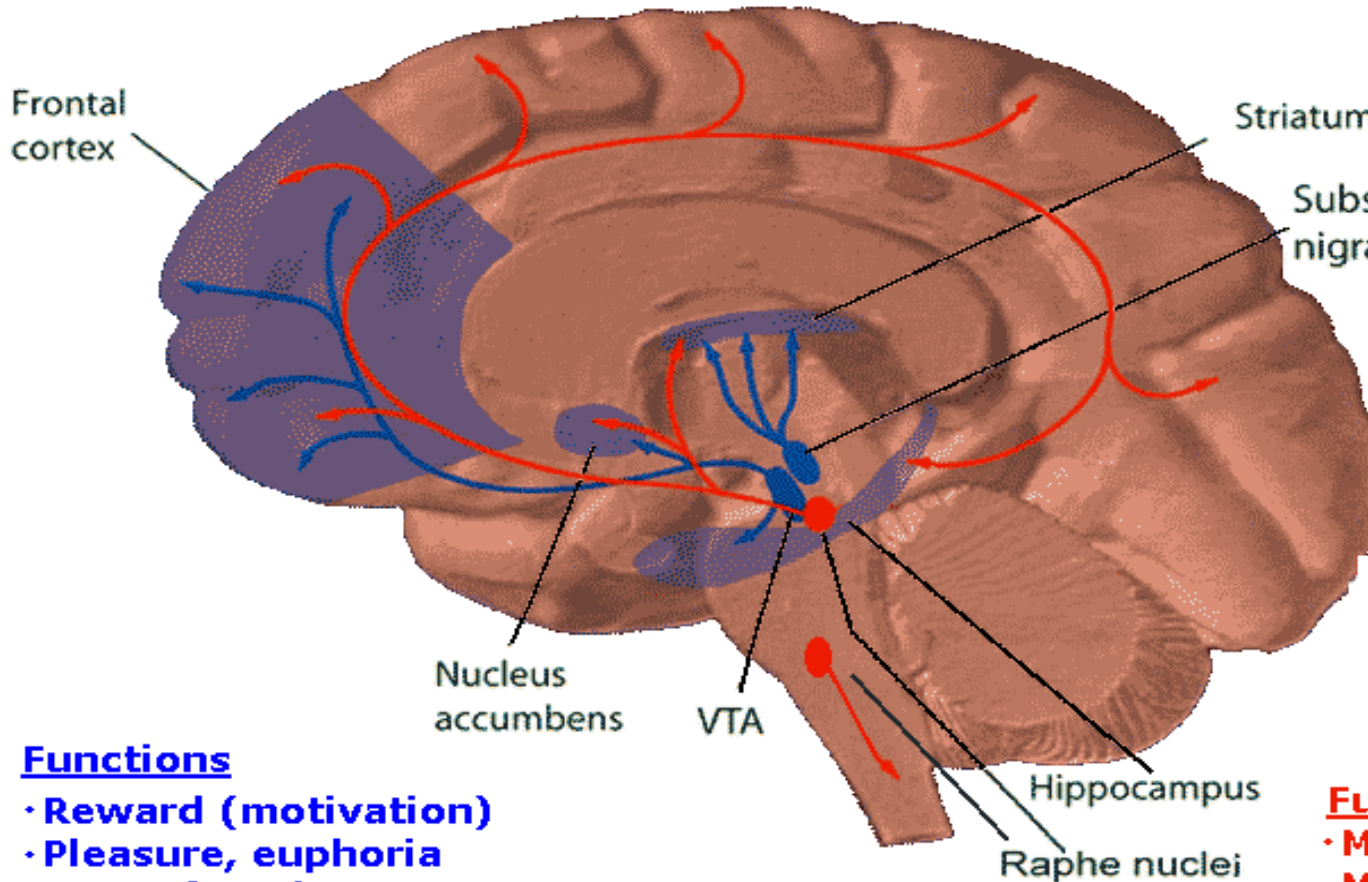
Двойка мъж и жена, жената се дразни от нещо дребно в мъжа, но не му казва и не го показва. Той не разбира, че прави нещо „грешно“ - в някои дни ходи на мач с приятелите си... Накрая жената къса с него и му казва „Ти не ме разбираш и не ме обичаш!“...

Биологични механизми

- „Пътека на награждаване (удоволствие)“ в мозъка (**Reward Pathway**)
- **Допамин** – успехът, изпълнението на цели, преодоляването, предусещането на успех, получаването на „наградата“ води до отделяне на допамин в челните дялове на кората. Допаминът усилва чувствителността и активността на невроните и подсилва нервните вериги, които са били активни непосредствено преди получаване на наградата.

Допаминов път

Dopamine Pathways

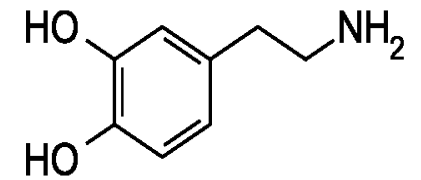


Functions

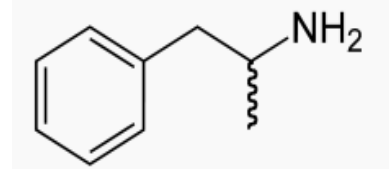
- Reward (motivation)
- Pleasure, euphoria
- Motor function (fine tuning)
- Compulsion
- Perseveration

Serotonin Pathways

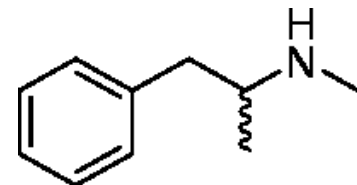
Dopamine



Amphetamine



Maetamphetamine



Functions

- Mood
- Memory processing
- Sleep
- Cognition

Хазарт, пристрастяване и допамин

- Шимпанзета, допамин – когато вероятността да получи награда е 100% или близко, нивото на допамин е няколко пъти по-ниско, отколкото ако шансът е 50/50.
- Пристрастяване към хазарт и компютърни игри.
- Мания и маниакално-депресивна психоза (циклофрения) - много високо ниво на допамин, предизвиква свръхактивност, хиперсексуалност, чувство за неуязвимост
- **Наркотици** – подобни на естествените невротрансмитери или променят работата на невроните.

Учене с подсилване

- **Универсален метод за самообучение в произволна среда.**
- Системата/агентът трябва да може да познае и измери кога получава награда и кога е наказан!

Двойка мъж и жена, жената се дразни от нещо дребно в мъжа, но не му казва и не го показва. Той не разбира, че прави нещо „грешно“ - в някои дни ходи на мач с приятелите си... Накрая жената къса с него и му казва „Ти не ме разбираш и не ме обичаш!“...

Любов...

Фрагменти

Надскача спънките житейски, дребни.
Изстрелваш се в вълшебен свят красив.
Поврежда се ракетата ти и те дебне
залязване на твоя път - до тук щастлив.

Животът става хубав, радостен, жадуван.
Едни очи достатъчни са за да си щастлив.
В смъртта се влюбваш, щом си ти разлюбен.
Прегръдка с нея търсиш, за да се спасиш.

Бездънен океан - зелено, синкаво, кафяво.
Безбрежен гладък хоризонт красив. Във него
скрита буря подло се задава. Прекършва
на тресчици корабът ти в риф бодлив.

...и т.н. ...

© „Фрагменти“, Т. Арнаудов, 2006

<http://bglog.net/blog/Tosh/site/posts/?bid=10445>

Любов

- **Пристрастяване** към естествените наркотици, отделяни от мозъка. „**Любимият**“ е стимул, който мозъкът запомня/приема като причиняващ отделянето на допамин, ендорфин, окситоцин и др. „Разбитото сърце“ е състояние на абстиненция, подобна на наркотичната. - Т. Арнаудов.
- *„Love is a stupid addiction.“* - Борис Казаченко
- Зацикляне в механизмите за учене с подсилване. - Т. Арнаудов.
- Анимизъм. - Т. Арнаудов.

Любов

- Локален максимум на функцията, търсеща удоволствие, около който разумът кръжи. - Т.А.
- **Вроден „бъг“ и основен недостатък на ученето с подсилване**, който се проявява още по-силно при наркотичната зависимост. Показва елементарните движещи сили на поведението на разумни системи, ако нямат допълнителни нива на управление и средства за извеждане на системата от „мъртва хватка“ (deadlock). - Т. А.

Какво е „щастие“?

- Нивата на допамин, серотонин, ендорфин, окситоцин и др. в различни части на мозъка...?

Учене с учител и подражание

- Ускорено обучение с по-малко налучкване.
- Учене на неща, които сами по себе си са неприятни и/или агентът не би ги правил (да ходиш на зъболекар, да ти бият инжекции, да пиеш горчиво лекарство, как точно да си попълниш данъчната декларация и пр. бюрокрации).
- Учене на неща, които са неутрални (факти).
- Мозъкът намира **изображение** между своите части на тялото и изглеждащите по подобен начин части на другите агенти/хора, записва възприятията за другите и ги възпроизвежда чрез своето тяло през изображението. - Т. Арнаудов

Учене с учител и подражание

- Разумът е йерархичен универсален подражател (симулатор) на въображаеми вселени. - Т. Арнаудов

Повече в лекцията за

Теорията на Разума и Вселената.

Защо са необходими?

- *Учене с подсилване* – универсален начин за приспособяване към всякаква среда, която може да задоволи основните нужди. Не е необходима външна намеса.
- *Учене с учител/подражание* – социални отношения и обществена йерархия, учене на факти, предвиждане на бъдещето чрез изпреварващо моделиране на възприятията.

Абстрактно за наградите и наказанията

- **Награда** – състояние, в което агентът/системата иска да се намира, стреми се да се намира там.
- **Наказание** – състояния, в което агентът/системата не иска да се намира, стреми се да не се намира там.
- **Неявни цели** на човека и на разумни машини.
- Повече – в следващите лекции.

Следва продължение

- Темата за ученето с подсилване и движещите сили на човешкото поведение **продължава** в:
 - Теорията за абстрактния универсален самообучаващ се интелигентен агент на Маркус Хутер
 - *Теорията на Разума и Вселената* на Т. Арнаудов

Връзки

-
-
-